



OPEN ACCESS

EDITED AND REVIEWED BY
Gavin T. L. Brown,
The University of Auckland, New Zealand

*CORRESPONDENCE

Xiaoming Zhai
✉ xiaoming.zhai@uga.edu

RECEIVED 19 May 2023

ACCEPTED 23 May 2023

PUBLISHED 14 June 2023

CITATION

Zhai X and Lu M (2023) Editorial: Machine learning applications in educational studies. *Front. Educ.* 8:1225802. doi: 10.3389/educ.2023.1225802

COPYRIGHT

© 2023 Zhai and Lu. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Editorial: Machine learning applications in educational studies

Xiaoming Zhai^{1,2*} and Min Lu³

¹AI for STEM Education Center, University of Georgia, Athens, GA, United States, ²Department of Mathematics, Science, and Social Studies Education, University of Georgia, Athens, GA, United States, ³Miller School of Medicine, University of Miami, Miami, FL, United States

KEYWORDS

artificial intelligence, machine learning, education, teaching, learning, assessment, student, deep learning

Editorial on the Research Topic

Machine learning applications in educational studies

Education, a cornerstone for human development, has witnessed significant transformations in recent years due to the advent of machine learning (ML) techniques. The confluence of big data, computational power, and increasingly sophisticated algorithms has enabled ML to make a substantial impact on the educational research and practice (Zhai et al., 2020b; Kubsch et al., 2022). As ML permeates various domains (e.g., Lu et al., 2021a,b; Lu and Liao, 2022), its potential to revolutionize educational research and practice becomes increasingly apparent (Zhai et al., 2020a; Zhai, 2021; Yürekli et al., 2022). A myriad of ML techniques, including supervised, unsupervised, and reinforcement learning, are employed to tackle diverse challenges and opportunities in the education (Zhai et al., 2020a, 2021). These applications span intelligent tutoring systems, adaptive learning platforms, learning analytics, and student performance prediction, among others (Linn et al., 2023).

In this Research Topic, we introduce a collection of articles that explore the wide-ranging applications of ML in educational studies, showcasing the power of these cutting-edge technologies to redefine the ways we assess, teach, and engage with learners. These contributions highlight the key methodologies, challenges, and opportunities associated with the implementation of ML in various educational contexts. Furthermore, they emphasize the ethical considerations, limitations, and challenges that must be addressed to ensure the responsible and equitable adoption of ML in education.

This paper, “Computer or teacher: who predicts dropout best,” (Eegdeman et al.) investigates the accuracy of teacher predictions of student dropout compared to ML algorithms and the potential benefits of combining both approaches. The research focuses on a vocational education program in sports with a relatively high dropout rate. It demonstrates that some teachers, as well as the composite of teacher predictions, could predict dropout more accurately than some ML algorithms at the beginning of the program. However, as the school year progresses and more data becomes available, ML algorithms’ performance improves, matching or surpassing teacher predictions. The study suggests that ML algorithms, when combined with teacher input, can lead to a more accurate and targeted approach to combatting student dropout. However, the research is limited by its small sample size, and further replication is needed for generalization. The results emphasize the potential benefits of leveraging both teacher insights and ML algorithms in dropout prevention programs and encourage future research on incorporating teacher predictions into ML algorithms to improve accuracy.

The study, “*Prediction of differential performance between advanced placement exam scores and class grades using machine learning*,” (Suzuki et al.) examines predictors of differential performance between AP class grades and AP exam scores using machine learning methods, specifically random forests, on data from 381 high school students enrolled in AP Statistics courses in the 2017–2018 academic year, and replicates the analysis on a separate cohort of 422 students from the 2018–2019 academic year. Results highlight students’ school and behavioral engagement as predictors of differential performance. The study suggests that high class grades don’t necessarily guarantee high AP exam scores, and school-level differences in relative performance pose equity concerns toward the use of AP exam scores in high-stakes decisions, such as college admissions. By identifying personal and contextual characteristics that predict performance, this research contributes valuable insights into the relationship between class grades and standardized test scores in the context of the AP program, which is prominent in US high schools and plays a significant role in educational decisions.

The study, “*Predicting attribution of letter writing performance in secondary school: a machine learning approach*,” (Boekaerts et al.) aims to investigate the factors that influence students’ causal ascription of their perceived writing outcome in vocational high schools using a machine learning approach. The study collected data from 1,130 students through prospective questionnaires and analyzed the interactions among domain-specific information, context-sensitive appraisals, and emotions to determine their impact on task engagement, task satisfaction, and attribution of the perceived learning outcome. The results indicate that the quality of the internal environment students create during the goal setting and striving stage influences their causal ascription of the perceived writing outcome. Additionally, motivational variables at both the domain and situation-specific levels significantly contribute to students’ attributions. The study’s findings suggest that attributions are more than individual difference variables and should be studied in a context-sensitive way. The study concludes that a system approach has the potential to bring together large bodies of research to better understand students’ attributions and that future research should consider the dynamic aspects of self-regulation and unfolding attribution processes.

Unlike the authors of the other three studies that utilized supervised learning approaches, Han et al. applied methods that are close to the unsupervised learning scope. Supervised learning approaches often specify at least one outcome in the model, whereas unsupervised learning lets the algorithm learn patterns in data without being explicitly told what the correct output should be. In other words, there is no specific outcome in the dataset. Some unsupervised learning algorithms share similarities with item response theory in that the goal is to measure the underlying factor of all the variables and maximize the variance explained by the model. Han et al. demonstrated that when the items in an exam are highly dimensional (for example, 300 items in their study), naively summing up the scores of all items could introduce noise/errors that reduce the quality of the measurement; on the other hand, utilizing item response theory or unsupervised learning

models could compress the items more efficiently. In this scenario, data-driven ML-based computational models and theory-based psychometrics are integrated into computational psychometrics.

Our aim with this Research Topic is to foster interdisciplinary dialogue and collaboration between the fields of machine learning and educational studies. By presenting a diverse array of perspectives and insights, we hope to inspire future research and advance the development of more effective, inclusive, and personalized educational experiences for learners worldwide.

As we navigate the complexities and potential of ML applications in education (Zhai et al., 2023), it is crucial to continue expanding our understanding, sharing best practices, and critically examining the implications of these technologies. We invite our readers to engage with the articles presented in this Research Topic, contributing to the ongoing conversation about the transformative potential of machine learning in educational studies (Zhai, 2021).

Author contributions

XZ drafted the manuscript. ML contributed to the study that she edited. XZ and ML read the manuscript. All authors contributed to the article and approved the submitted version.

Funding

The study was funded by National Science Foundation (NSF) (Award \# 2101104, 2138854, PI: XZ). This study was also funded by the Department of Public Health Sciences 2023 Copeland Foundation Project Initiative Award, University of Miami (to ML).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Author disclaimer

Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the NSF.

References

- Kubsch, M., Krist, C., and Rosenberg, J. M. (2022). Distributing epistemic functions and tasks—A framework for augmenting human analytic power with machine learning in science education research. *J. Res. Sci. Teach.* doi: 10.31219/osf.io/sg9jk
- Linn, M. C., Donnelly-Hermosillo, D., and Gerard, L. (2023). “Synergies between learning technologies and learning sciences: promoting equitable secondary school teaching,” in *Handbook of Research on Science Education* (Routledge), 447–498. doi: 10.4324/9780367855758-19
- Lu, M., and Liao, X. (2022). Access to care through telehealth among U.S. Medicare beneficiaries in the wake of the COVID-19 pandemic. *Front. Public Health* 10, 946944. doi: 10.3389/fpubh.2022.946944
- Lu, M., Parel, J.-M., and Miller, D. (2021a). Interactions between staphylococcal enterotoxins A and D and superantigen-like proteins 1 and 5 for predicting methicillin and multidrug resistance profiles among *Staphylococcus aureus* ocular isolates. *PLoS ONE* 16, e0254519. doi: 10.1371/journal.pone.0254519
- Lu, M., Sha, Y., Silva, T. C., Colaprico, A., Sun, X., Ban, Y., et al. (2021b). LR hunting: a random forest based cell–cell interaction discovery method for single-cell gene expression data. *Front. Genet.* 12, 708835. doi: 10.3389/fgene.2021.708835
- Yürekli, H., Yigit, Ö. E., Bulut, O., Lu, M., and Öz, E. (2022). exploring factors that affected student well-being during the COVID-19 pandemic: a comparison of data-mining approaches. *Int. J. Environ. Res. Public Health* 19, 11267. doi: 10.3390/ijerph191811267
- Zhai, X. (2021). Practices and theories: How can machine learning assist in innovative assessment practices in science education. *J. Sci. Educ. Technol.* 30, 139–149. doi: 10.1007/s10956-021-09901-8
- Zhai, X., Haudek, K. C., Shi, L., Nehm, R., and Urban-Lurain, M. (2020a). From substitution to redefinition: a framework of machine learning-based science assessment. *J. Res. Sci. Teach.* 57, 1430–1459. doi: 10.1002/tea.21658
- Zhai, X., Neumann, K., and Krajcik, J. (2023). AI for tackling STEM education challenges. *Front. Educ.* 8, 1183030. doi: 10.3389/feduc.2023.1183030
- Zhai, X., Shi, L., and Nehm, R. (2021). A meta-analysis of machine learning-based science assessments: factors impacting machine-human score agreements. *J. Sci. Educ. Technol.* 30, 361–379. doi: 10.1007/s10956-020-09875-z
- Zhai, X., Yin, Y., Pellegrino, J. W., Haudek, K. C., and Shi, L. (2020b). Applying machine learning in science assessment: a systematic review. *Stud. Sci. Educ.* 56, 111–151. doi: 10.1080/03057267.2020.1735757